

Advancing the Evidence Agenda at USAID

September 26, 2017

Sarah Rose and Amanda Glassman



Photo credit: Staff Sgt. Jeffrey D. Anderson/U.S. Marine Corps

The authors would like to thank Masood Ahmed, Erin Collinson, Michael Eddy, Cindy Huang, Ruth Levine, Scott Morris, and other anonymous reviewers for helpful conversations and comments.

Introduction

Front and center in discussions around the reform and redesign of the United States Agency for International Development (USAID) are the objectives of increased efficiency and effectiveness. The agency's new administrator, Mark Green, who has highlighted these goals from day one, has an excellent opportunity to improve the agency's efficiency and effectiveness through better generation and use of evidence to inform policy and programming decisions.^[1]

USAID talks a lot about its results, but much of what it highlights are outputs.^[2] Much less is said about actual outcomes or development impact, even though understanding these is critical for developing effective programs that are worth the money put into them.

USAID has improved its generation and use of evidence in recent years, but several factors continue to constrain the agency's evidence orientation. If Administrator Green is serious about enhancing USAID's efficiency and effectiveness, it is imperative that he take steps to address these barriers and ensure the routine use of evidence as part of US development policy and programming. In the absence of this, the effectiveness of any agency redesign effort will ultimately fall short of its potential.

This note offers eight suggestions to advance evidence generation and use at USAID; they are described in detail below.

1. Elevate and consolidate the evidence agenda with the establishment of a new unit: Evidence, Evaluation, and Learning (EEL)
2. Create incentives for improved evaluation quality through a public scoring system
3. Focus on synthesizing data for greater accessibility and use
4. Streamline reporting requirements to allow greater focus on more useful evidence
5. Evaluate staff performance on evidence use
6. Adjust staffing choices and opportunities to better emphasize skills in evaluation
7. Bake evidence-based programming into the procurement process
8. Continue to support and invest in external organizations' contributions to learning

A base upon which to build

Over the last decade, USAID has taken important steps toward becoming a more evidence-oriented agency. Its well-regarded 2011 Evaluation Policy paved the way for more evaluations and helped focus the agency on improving their quality, timeliness, and utility. These evaluations are being used to inform decision making, too. A 2016 study found that 90 percent of the evaluations reviewed were reported by USAID staff to have influenced some kind of programmatic action.^[3]

The launch of Development Innovation Ventures (DIV) in 2010 was another advance in evidence-oriented learning. The main function of DIV is to identify and rigorously test new solutions to development problems and help scale those that prove successful. It has pushed the envelope on development evidence in other ways, too. For instance, DIV is collaborating with other donors to design an outcomes-based payments fund, including the establishment of a Development Impact Bond.^[4] It is also pioneering an effort to develop cost-effectiveness “hurdle rates” against which to benchmark traditional non-cash-based USAID programs. This framework could serve as a basis for a more institution-wide assessment of the cost-effectiveness of future proposed aid programs, something current USAID programming lacks in any consistent and rigorous way.

Not all of USAID’s evidence-oriented work is recent, of course. For instance, USAID has long sponsored the generation of high-quality data as a public good through its support for the

Demographic and Health Surveys program. These surveys have long provided a consistent cross-sectional snapshot of global health, fertility, mortality, and socioeconomic data, and are one of the first household surveys conducted in many low-income countries. Their data is widely used for evaluation and learning purposes, though their use can be enhanced.

Nine barriers to improved use of evidence and evaluation

Despite progress, USAID continues to face challenges that limit its ability to use evidence to inform more effective programming and impede greater contributions to the global body of evidence. These challenges fall into nine main categories:

- **Availability constraints.** The state of evidence in global development has improved over the last 10 years, but our understanding of the effectiveness or net impact of most development programs still falls short.^[5] Accordingly, it is hard to use evidence to inform programming when the evidence simply does not exist. Though USAID has contributed to filling the gap, both with its own evaluations and through support for external evaluation institutions, it is not adequately capturing opportunities to understand the impact of its activities. Most of its evaluations focus on accountability questions (was money spent as intended? was the intervention implemented as planned?) rather than learning questions (did the intervention achieve its objective? what kinds of interventions achieve the outcome of interest?). While it is wrong-headed to blindly push for impact evaluations rather than seeking the right approach for the question to be answered, the extremely small proportion of USAID impact evaluations (less than 1 percent, according to two recent studies) suggests that the agency is overlooking opportunities to answer questions about what works.^[6] Part of what constrains USAID from pursuing more rigorous evaluations is that much of what the agency finances is technical assistance and training (TAT)—paying nongovernmental organizations or firms to provide advice or training to governments or community organizations so that these can, in turn, implement programs and policy more effectively. Here the challenge is compounded. Not only is the evidence for the effectiveness of the supported program or policy often unavailable, but donors also tend to have a hard time directly attributing any outcomes achieved to TAT. Though there are good examples of rigorous evaluations of TAT interventions, at USAID, they are scarce.^[7]
- **Accessibility constraints.** Even where evidence does exist, it is not always *accessible*. That is, main findings about a project or across a sector are often not synthesized and presented

in a digestible manner for very busy staff, some of whom have little training on concepts relevant to consuming evidence. Even when synthesized information is available, it is not always made available within a useful timeframe for decision making.^[8] Though USAID has developed and promoted several tools and processes to help missions learn from evaluation findings, these are not always widely used.^[9]

- **Quality constraints.** The quality of USAID evaluations has improved since the release of the evaluation policy, but two recent studies found that a material percentage of USAID’s evaluations fail to meet basic standards of quality.^[10] Many evaluations did not use appropriate study design, sampling, data collection, and/or analysis, and often the evidence did not support the findings and recommendations. The latter is particularly concerning since staff seeking to use the results of an evaluation to inform programming are likely to turn directly to the recommendations rather than spend time assessing their reliability given the methodology. Quality problems are sometimes the result of underperforming evaluation partners. They can also stem from inappropriate technical direction to the evaluation partner from USAID staff inexperienced in evaluation.
- **Time constraints.** USAID staff are often under intense pressure to execute programming, manage contracts, and feed the proverbial “beast” with burdensome reporting requirements. As a result, they have less time to seek relevant evidence and apply its lessons to current programming. Engaging in creative thinking about how to use USAID programming to fill evidence gaps also tends to fall down the priority list. In addition, plans for rigorous evaluations that would add to the body of evidence can sometimes be scuppered when processes required for the evaluation would end up slowing implementation.
- **Structural constraints.** Changing organizational culture to embrace evidence is a long-term process that needs a high-level champion. Unfortunately, USAID has rarely had an empowered person with a strong institutional voice promoting evidence in a sufficiently cross-cutting position to influence change throughout the organization. The Bureau for Policy Planning and Learning (PPL) is the headquarters of the agency’s learning, evaluation, and research (LER) efforts, but it has limited reach to missions where most evaluation efforts are managed. Separately, DIV supports high-quality, rigorous evaluation to test solutions to development challenges. However, its small size and limited connection with missions gets it even less attention. Furthermore, because both LER and DIV are components of larger bureaus—PPL for LER and the Global Development Lab (the Lab) for DIV—they and their focus on evaluation and evidence can be overshadowed by their respective bureaus’ other activities.

- **Skills constraints.** USAID is a huge, decentralized organization. While LER has valuable evaluation resources for missions, the limited number of LER staff simply cannot provide deep support everywhere USAID works. The individual bureaus also maintain monitoring and evaluation (M&E) advisors, and while these are structurally better placed to reach mission staff, they are also few in number and similarly unable to provide deep support across the board. Furthermore, most M&E staff have generally focused more on traditional approaches to M&E than pushing the envelope on new approaches to evidence-generation.
- **Funding constraints.** Funds to USAID’s evidence engines are under threat. The Trump administration’s FY2018 budget request included a cut of 44 percent to PPL and 85 percent to the Lab compared to FY2016.^[11] While the proposed budget does not necessarily reflect final FY2018 appropriation levels, the uncertainty around cuts appears to be putting DIV, the Lab’s evidence-focused program, on the chopping block. It announced in July its “temporary” suspension of applications for new awards.
- **Incentive constraints.** USAID staff have poor incentives to make an impact evaluation work. New activities can take several years to design and procure, during which time, an impact evaluation should also be considered and set up. Because of the rotation cycle of American staff, however, those in charge of design will often not see the activity’s implementation, and those overseeing implementation were often uninvolved in design. There is little incentive, therefore, for a foreign service officer to prioritize evaluation at the design stage when s/he will not be present when findings are released. And when implementation is complicated by evaluation, foreign service officers, whose primary role is managing program agreements, will tend to make decisions that favor implementation rather than a rigorous evaluation.
- **Existential constraints.** In the current environment, US foreign assistance—and especially USAID—is facing somewhat existential questions: What will be the magnitude of future budget cuts? Will missions close? Will USAID functions get folded into State? In such a context, evidence and learning (which must focus on failures as well as successes) can be downplayed in favor of advocacy for the overall aid enterprise.^[12] This tooth-and-nail fight to preserve the structures and budgets of US foreign assistance, can reinforce the perceived need to avoid suggestion or evidence that aid programs are not working perfectly well. The irony, of course, is that suppressing frank acknowledgement of failure is at odds with one of the objectives of those who would seek to cut or restructure foreign assistance; honest discussion of evidence as part of learning is necessary for greater efficiency.

Eight ideas to address constraints to the use and generation of evidence

It now falls to Administrator Green to champion both new and ongoing efforts to generate and use more and better evidence in pursuit of more effective and efficient aid programs. Here are eight ideas for implementation:

1) Elevate and consolidate the evidence agenda with the establishment of a new unit: Evidence, Evaluation, and Learning (EEL)

EEL would consolidate and expand the evidence, evaluation, and learning functions currently housed in PPL and the Lab, bringing together LER and evidence-focused pockets of the Lab, such as DIV. While the core work of LER and DIV are different—LER supports the implementation of the agency’s evaluation policy and DIV competitively awards grants to those with innovative ideas to address development challenges and then rigorously tests them to see what works—both are focused on building the evidence base around what development interventions work and transferring that knowledge to program staff who can design and implement programs accordingly. Establishing a combined independent unit would better highlight and bring to the fore the importance of this work, rather than subsuming it under multiple larger bureaus. It could also strengthen and streamline the two units’ overlapping roles, like capacity building around evidence and dissemination of evidence-based learning. The joint unit would support Administrator Green’s efforts to understand, manage, and advocate for what works in the use of US development aid, and contribute to the agency’s reporting to Congress on impact and cost effectiveness.

The head of EEL should report directly to the administrator.^[13] The reporting arrangement would also create high-level champions for evidence, something the agency currently lacks but which is critical for institutionalizing evidence generation and use. Changing the way an organization works is difficult, and efforts to do so can fall by the wayside without strong political support. On the other hand, history suggests that an administrator who demonstrates routine interest in the use of rigorous evaluation results and other evidence can have significant influence. Observers point to Administrator Douglas Bennet, 1979-1981, and his encouragement of systematic and rigorous evaluation as an early high point in the agency’s evaluation history.^[14]

Beyond consolidating the current functions of LER and DIV, EEL should expand efforts to identify needs and opportunities for evidence generation and use. One option is to work with missions to explore ways to lower the cost of evidence generation, whether through innovative uses of technology (e.g., satellite data, cell phone records) or other mechanisms. This is especially

important in fragile and conflict-affected areas, where costs of evidence collection are often higher, situations are fluid, and findings tend to be less generalizable to other places.

Another option is to provide long-term (several month) embedded support to help overextended mission staff think through opportunities to generate evidence, and work with policymakers and counterparts to determine what questions an evaluation should try to answer.^[15] EEL would identify major evidence gaps and strive to make new contributions to learning where the evidence base is weakest and where there is significant possibility to inform critical future decisions that the agency (and its partners) will have to make (e.g., how to improve learning for children in conflict areas).^[16] The unit would also be charged with synthesizing and communicating research results at a sector or subsector level. A similar model has been used with success by the General Services Administration's Office of Evaluation Sciences (formerly the White House Social and Behavioral Sciences team), which has collaborated with USAID's Global Health Bureau to provide this type of support to missions. Similarly, the World Bank's Africa Gender Innovation Lab supports the World Bank and other donors, including USAID, in building rigorous evidence generation and use into programming. In fact, Gender Innovation Lab staff calculated that for each dollar they spend, 46 project dollars are directly influenced.^[17]

2) Create incentives for improved evaluation quality through a public scoring system

Three recent studies assessed a sample of USAID evaluations against specific quality criteria, but individual evaluation reports are not routinely scored for quality. USAID should adopt standard quality criteria and score each evaluation report based on how well it meets the specified standards—a potential function for EEL. Recent CGD work proposes a framework for quality assessment that looks at the relevance of objectives, relevance of data, sampling validity and reliability, and analytical validity and reliability.^[18] Establishing a set of quality criteria would communicate expectations clearly to evaluation partners and help USAID M&E staff manage evaluation contracts for quality. Taking this a step further by posting the score along with the evaluation report would serve a two-fold purpose. It would provide incentives both for evaluation partners to produce high quality evaluations and for USAID staff managing evaluation contracts to provide appropriate technical guidance. In addition, a clear indication of quality could help potential evaluation users make more informed decisions about how to interpret results.

3) Focus on synthesizing data for greater accessibility and use

As the Trump administration's FY18 Congressional Budget Justification for the Department of State, Foreign Operations, and Related Programs states, "The true value of data analysis, performance monitoring, and program evaluation is only realized if the lessons they reveal are used to inform and support foreign assistance programs and projects."^[19] USAID's record in this area is mixed. Most evaluations have informed program design or implementation decisions to

some degree, according to a survey of USAID staff. However, the same study found the agency could do more to encourage greater evaluation use.^[20]

Evaluation use hinges in large part on how findings are disseminated. The minimum requirement for disseminating USAID's evaluation findings is to post reports to the Development Experience Clearinghouse (DEC), the agency's online portal that houses evaluations alongside a variety of other documents. This, however, is insufficient for a number of reasons. It is extremely difficult to find relevant documents on the cumbersome site, and those seeking evidence rarely have the time and motivation to try to locate and read through multiple long reports. The DEC's value as a repository for evidence is also limited in that it only collects USAID's own evaluations, not capturing a wide variety of useful evidence from other US agencies like the Millennium Challenge Corporation, or external outfits like the World Bank, the International Initiative for Impact Evaluation (3ie), and the Abdul Latif Jameel Poverty Action Lab.

USAID recognizes that publication alone is insufficient for generating use and has taken important steps to bolster dissemination of evidence through efforts like establishing internal communities of practice at the bureau level and creating technical-level linkages between evaluation specialists and sector or mission teams. Still, there is room for further improvement. A study on evaluation utilization found that mission staff often reported being unaware of evaluation findings in their sector from other missions (and, presumably, from non-USAID sources as well) or having difficulty accessing this kind of information. In fact, better synthesis and dissemination of evidence and evaluation findings at the sector level was widely requested by interviewed missions, a finding USAID has begun to address.^[21] The agency under Administrator Green should continue efforts to devote more staff resources to synthesizing evidence by sector/subsector—with due attention to questions of generalizability across different environments—and ensure communication of synopses with posts on a regular basis.^[22]

4) Streamline reporting requirements to allow greater focus on more useful evidence

USAID should cooperate with Congress, where relevant, to streamline the agency's many reporting requirements. USAID's M&E staff are, as their title suggests, responsible for both monitoring (collecting data for specific indicators to show how a project is progressing and whether objectives are being achieved) and evaluation. The problem is that M&E staff often spend so much time compiling and reporting monitoring data for a wide range of mandated reports, that they have limited time to focus on other critical tasks. Some of the essential tasks that may fall by the wayside are ensuring good data quality for the measures most relevant for project management and working on evaluations—identifying opportunities to evaluate, developing high-quality scopes of work, monitoring the work of evaluators for quality, participating on evaluation teams, and disseminating evaluation findings.

5) Evaluate staff performance on evidence use

It is the responsibility of USAID managers to clarify how staff are expected to include evidence in program design and management. The current Foreign Service Skills Matrix includes only vague references to applying lessons learned.^[23] A more explicit assessment of staff competence in this area should be included in performance reviews and promotion criteria.

6) Adjust staffing choices and opportunities to better emphasize skills in evaluation and evidence use

Knowing that a significant portion of an M&E officer's time will be consumed with reporting requirements, evaluation skills have sometimes been undervalued as a relevant qualification. At the mission level, in particular, M&E staff should be hired according to their familiarity and experience with rigorous quantitative and qualitative evaluation methods. While USAID offers M&E field staff core training in evaluation policy and practice, it is insufficient to build the necessary expertise to design and manage a high-quality evaluation. USAID has added and will continue to expand upon new specialized evaluation training modules, but classroom training (especially, as is typical, in the absence of a subsequent proficiency test) will not substitute for demonstrated skills and experience.

While evaluation trainings (which have now been offered to over 1,600 USAID staff) will not, in and of themselves, create a new cadre of experienced evaluation professionals, the agency should continue to encourage them for all staff. They serve an important role in familiarizing staff with the evaluation process, and, over the medium term, helping them internalize that part of their job is the systematic application of learning from evidence.

Beyond trainings, mission M&E staff should take better advantage of hands-on learning opportunities, whether through the M&E fellows program or through greater participation on evaluation teams. A report on evaluation usage at USAID found that participation in an evaluation instilled a much better understanding of the evaluation process, as well as more confidence in the ability to manage evaluation products for quality.^[24]

USAID should also consider separating monitoring and evaluation functions into two staff roles at the mission level. Though monitoring data can form an important part of many evaluations, monitoring and evaluation are distinct disciplines. Freeing evaluation-focused personnel from spending much of their time on reporting requirements could allow USAID to attract and retain individuals better qualified in evaluation design and management.

7) Bake evidence-based programming into the procurement process

Administrator Green has expressed interest in procurement reform at USAID. As he seeks to reform current procedures, the administrator should consider how to use the procurement process to better ensure that proposals are crafted to reflect the current state of evidence and/or build in opportunities to generate evidence. He should also encourage more creative use of existing mechanisms to support results-based funding.

One way to do this would be to include in requests for application a scoreable requirement to assess how well a bidder demonstrates an understanding of the existing evidence base relevant to the project and incorporates economic evaluation and evidence into their proposal. Such a requirement would have to be developed in close collaboration with contracting officers, who tend to prefer unambiguous specifications rather than those (like whether a firm demonstrates capability to use evidence well) that require more judgement.

A related requirement should be for the USAID staff member writing a request for proposal (RFP) to be responsible for citing the latest research in the sector. This creates an understanding that the implementer needs to know the literature and build upon it for their proposal to be successful. USAID's ability to ground its RFPs in evidence is especially important for the agency's acquisitions processes since these award types give USAID substantial technical control over project design and implementation; the implementing partner must do exactly as the contract specifies, even if available or new evidence suggests better practices. Though a minority of USAID procurements are contracts, one could argue that until and unless USAID is better able to incorporate a strong understanding of the relevant evidence in its RFPs and in its subsequent technical guidance to awardees, it should seek to shift even more toward greater use of assistance awards (e.g., grants, cooperative agreements, fixed amount awards) that allow USAID staff and implementing partners to more easily identify and implement adjustments based on evidence that emerges during implementation.

A unit like the proposed EEL could have a role in ensuring that contracts, especially large contracts, are sufficiently based on evidence. It could also be charged with working with the Office of Assistance and Acquisition to sort out how existing procurement mechanisms could be used more creatively to focus on paying for results.

8) Continue to support and invest in external organizations' contributions to learning

USAID is an active member, participant, and supporter of several external organizations' efforts to generate and use evidence. The agency sits on the board of 3ie; it launched and supports the Global Innovation Fund (GIF), which uses a DIV-like pilot-test-and-scale approach to solving

development challenges; and it funds researchers and evaluation experts to hone and broaden approaches and methods for evaluation.

However, continued support for some of these initiatives is under threat. USAID lags in fulfilling its commitment to GIF, for example. Unlike DIV, GIF is a source of both grant and risk capital (debt and equity) investments to innovate to reduce poverty, rigorously measure results, and scale up what works—a perfect fit with the stated objectives of the new administrator with its strengthened focus on crowding in private funding. With the future of the Global Development Lab uncertain, however, USAID’s continued support of GIF is also unclear.

Withdrawing support from GIF and other external evidence-focused initiatives would be misguided. Evaluation serves two functions: (1) day-to-day accountability, to see if inputs were deployed as intended; and (2) accountability for outcomes and learning, to create evidence to inform better program design in the future. Support for external efforts serves the second function. Learning can and should come from many more sources than just the evaluation of USAID’s own programs, especially when so many of USAID’s own evaluations are performance evaluations more suited for routine tracking than for the more profound and important accountabilities related to outcomes and learning from what does and does not work.

Conclusion

As a reorganization of USAID is considered, data, evidence, evaluation, and learning should assume a central role. If monies are reallocated to new priorities, such as an expansion of malaria elimination efforts, how will the administration show that the monies are optimally allocated for impact and that they have—indeed—contributed to malaria elimination? If country graduations are a major emphasis, what strategies can be shown to work best to create incentives for greater partner country investments in shared priorities, and for sustained impact against goals?

As the administrator starts his mandate, he may consider how he will be able to describe his impact as his tenure is completed—will he talk about numbers of people trained, or amounts of products purchased? Or will he be able to describe the attributable difference that USAID made to the people in need that the agency is committed to helping? The latter narrative will only be possible with a renewed commitment to evidence, evaluation, and learning at the highest levels.

[1] Administrator Green highlighted a focus on efficiency and effectiveness in his welcome remarks to USAID staff. Green, Mark. "USAID Administrator Mark Green Welcome Remarks to Employees." August 7, 2017. United States Agency for International Development. A memo from the Office of Management and Budget instructed agencies to demonstrate how they would "build and use a portfolio of evidence to improve effectiveness." OMB Memorandum M-17-22 (April 12, 2017). *Comprehensive Plan for Reforming the Federal Government and Reducing the Federal Civilian Workforce*. p. 10.

[2] For instance, USAID's 2011-2015 Education Strategy Progress Report summarizes progress in terms of things like number of textbooks provided and number of teachers trained rather than any measure of actual learning outcomes.

[3] Hageboeck, Molly, Micah Frumkin, Jenna L. Heavenrich, Lala Kasimova, Melvin Mark, and Aníbal Pérez-Liñán. 2016. *Evaluation Utilization at USAID*. Arlington: Management Systems International.

[4] For more on Development Impact Bonds, see Center for Global Development and Social Finance. 2013. *Investing in Social Outcomes: Development Impact Bonds, The Report of the Development Impact Bond Working Group*. Washington, DC: Center for Global Development.

[5] Rose, Sarah. 2017. *Some Answers to the Perpetual Question: Does US Foreign Aid Work—and How Should the US Government Move Forward with What We Know?* Washington, DC: Center for Global Development.

[6] Hageboeck et al., 2016; Hageboeck, Molly, Micah Frumkin, and Stephanie Monschein. 2013. *Meta-Evaluation of Quality and Coverage of USAID Evaluations 2009-2012*. Arlington: Management Systems International.

[7] Ruprah, Inder, and Luis Marcano. 2009. Does Technical Assistance Matter? An Impact Evaluation Approach to Estimate its Value Added. *Journal of Development Effectiveness* 1 (4): 507-528; Dunsch, Felipe, David Evans, Ezinne Eze-Ajoku, and Mario Macis. 2017. *Management, Supervision, and Health Care: A Field Experiment*. NBER Working Paper No. 23749. Cambridge: National Bureau of Economic Research.

[8] Hageboeck et al., 2016.

[9] Hageboeck et al., 2016.

[10] Hageboeck, Frumkin, and Monschein, 2013; Goldberg Raifman, Julia, Felix Lam, Janeen Madan Keller, Alexander Radunsky, and William Savedoff. 2017. *Evaluating Evaluations: Assessing the Quality of Aid Agency Evaluations in Global Health*. CGD Working Paper 461. Washington, DC:

Center for Global Development; United States Government Accountability Office. 2017. *Foreign Assistance: Agencies Can Improve the Quality and Dissemination of Program Evaluations*. Washington, DC: GAO. According to the GAO study, USAID performed near the middle of the six agencies assessed in terms of the distribution of evaluations across categorizations of high, medium, and low quality.

[11] Department of State. 2017. Congressional Budget Justification. Department of State, Foreign Operations, and Related Programs, Fiscal Year 2018.

[12] Morris, Scott. 2017. The Damage Already Done to the Foreign Assistance Budget. Center for Global Development. *US Development Policy Blog*.

[13] The head could fulfill the functions of a chief evaluation officer, a position recommended for all government agencies by the congressionally established Commission on Evidence-Based Policymaking to lead the full range of evidence-related activities. Commission on Evidence-Based Policymaking. 2017. *The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking*.

[14] Butterfield, Samuel. 2004. *U.S. Development Aid--An Historic First: Achievements and Failures in the Twentieth Century. Contributions to the Study of World History* (Book 108). Westport: Praeger.

[15] The existing Monitoring and Evaluation Fellows program, which embeds M&E experts in missions for six months to two years, does this to an extent. However, fellows' efforts tend to be more focused on building staff capacity for the standard types of program evaluation USAID typically does rather than exploring more social and behavioral sciences questions.

[16] Papadopoulos, Nina and Meghan Mattern. 2017. What Does 'Back to School' Mean for Children in Crisis and Conflict? United States Agency for International Development. *Impact Blog*.

[17] Buvinic, Mayra and Rebecca Furst-Nichols. 2015. World Bank Africa Gender Innovation Lab (GIL) Mid-term Review.

[18] Goldberg Raifman et al., 2017.

[19] Department of State, 2017. P. 11.

[20] Hageboeck et al., 2016.

[21] Hageboeck et al., 2016.

[22] The Commission on Evidence-Based Policymaking discusses the importance of agencies having “knowledge brokers” to serve as intermediaries between producers and users of evidence. (Commission on Evidence-Based Policymaking, 2017.)

[23] United States Agency for International Development. 2015. *AEF Foreign Service Skills Matrix*. AID 461-4.

[24] Hageboeck et al., 2016.